

# Shaping geometric objects by cumulative translational sweeps

by R. C. Evans  
G. Koppelman  
V. T. Rajan

**This paper introduces the cumulative translational sweep (CTS) as a tool for shaping geometric objects. It describes how it may be applied, in combination with Boolean operations, to stimulate growth and shrinking over the boundary regions of polyhedral models, and how, by creating additional facets, it may be used to achieve global rounding effects along model edges and around their vertices. CTSS are examined in terms of a conceptual framework that describes their effects as Minkowski sums—of the polyhedra to be swept, with convex polyhedra from the class of mathematical objects known as zonotopes. Included is a discussion of applications in the OYSTER program, a CAD system for the simulation of semiconductor wafer fabrication.**

## Introduction

Techniques described in this paper introduce a new method—and its first disciplined use—for the shaping of polyhedral models. It is based upon the theoretical construct

of a cumulative translational sweep (CTS), and its application achieves shaping effects that are interpretable in terms of Minkowski sums.

Our interest in this subject reflects an effort to enhance the capabilities of the OYSTER system [1], which strives to realistically represent device fabrication process effects in terms of polyhedra of the GDP modeler [2, 3]. We have considered some shaping techniques alternative to the one described here, including general offsetting techniques leading beyond the polyhedral domain [4–6], polyhedral offsetting applied to convex polyhedra [7], and a method for splitting general polyhedra into convex components that are separately shaped and reassembled [8]. The method we present is applicable to general polyhedra, makes no demands for partition and reassembly, and generates polyhedral results.

“Shaping” has many senses, such as growing, shrinking, rounding, filleting, faceting, blending, and smoothing. It is a generic term in geometric modeling, encountered in such applications as growing and shrinking to solve the collision-avoidance problem [9]; growing and shrinking for the generation of blends [4, 5]; sweeping to compute the shape of various space regions [10, 11]; and offsetting as a means of defining mechanical tolerance.<sup>1</sup>

Sweeping, as a geometric modeling tool, refers in its broad sense to the tracking of a body's motion in space. Most

©Copyright 1987 by International Business Machines Corporation. Copying in printed form for private use is permitted without payment of royalty provided that (1) each reproduction is done without alteration and (2) the *Journal* reference and IBM copyright notice are included on the first page. The title and abstract, but no other portions, of this paper may be copied or distributed royalty free without further permission by computer-based and other information-service systems. Permission to *republish* any other portion of this paper must be obtained from the Editor.

<sup>1</sup> V. Srinivasan and R. Jayaraman, IBM Research Division, Yorktown Heights, NY; private communications.

modelers can compute the swept volume, or sweptspace, of a moving body which does not tumble, i.e., which has only translational freedom; this is the region of space through which the moving body has passed, or, quite informally, its "ghost." Some modelers, such as the GDP modeler supporting OYSTER, can approximate tumbling motions (rotational freedom) as well.<sup>2</sup> No modelers we know of construe a sweeping operation as effecting motion not only in the body but also in its ghost; however, if the ghost is deemed to be carried along with the body and so to spawn its own ghost of a ghost (and so on, and so forth, . . .), then there is effected a much larger sweptspace than is usually conceived—one that we call cumulative to contrast it with the familiar tubular sweptspace. Cumulative sweeping without tumbling serves our shaping needs; as noted above, we call such an operation a cumulative translational sweep, or CTS. We interpret motion curves as prescriptions for CTS shaping operations that may be performed upon an input body, interpreting sweptspaces as shaped output bodies.

The CTS method may be related to a general offsetting operation known as Minkowski summation—and also as set convolution [12]—that shapes a set of points by adding (vector summing) to each of them, in all possible pairings, each of the points of another selected shaping body. The effect of CTS prescriptions that we apply in our work is to develop around any input polyhedron a polyhedral sweptspace which is the Minkowski sum of the input with a shaping polyhedron from the polytope subfamily of zonotopes [13–16].

Serra [17] develops the algebraic properties of "dilation" and "erosion"—shaping operations based on neighborhood rules applicable over discrete domains (such as the arrangement of pixels in an image). Some sequences of these operations are expressible as Minkowski sums and were implemented in image processing architectures [18].

In the first section of the paper we describe various sweep types, using a classification scheme that includes the CTS; we then describe the particular CTS type developed for OYSTER; sections follow that give theoretical justifications and present some examples of zonotope shaping bodies; there follows a discussion of CTS applications in OYSTER; and, in a final section, we consider a theory for the general CTS.

### Sweep types in general

A *sweep* may be generally defined as any function  $S$  that maps a time  $t \in [0, T]$ , a motion  $F(t)$ , and a body  $B$  to a *sweptspace*  $S(F, t, B)$  that depends on both  $B$  and its motion history. To avoid encumbering notation, we often suppress explicit reference to  $F$  when referring to the sweptspace and use  $S(t, B)$  as its designator.

Important uses of the term "sweep" (within the topic of generalized cylinders) admit motions  $F$  that involve scaling. For our purposes,  $F$  is not allowed such freedom but is constrained to be a rigid motion. It is taken as a time-parameterized frame transformation having translation vector  $\mathbf{f}(t)$  as its *translational component* and rotation matrix  $M(t)$  as its *rotational component*. It associates with any point,  $\mathbf{b}$ , a *trajectory curve*,  $\mathbf{f}(t) + M(t)\mathbf{b}$ . We assume that  $M(0)$  is the identity frame and that both  $\mathbf{f}$  and  $M$  are continuous and piecewise-differentiable.

Various sweep types are definable in terms of motion  $F$  and the rules that determine sweptspace membership. After making some broad distinctions on these bases, we will consider all rigid motion sweeps to be either tubular or cumulative.

The *trajectory*, or *tube*, of  $B$  under sweep  $S$  is the union of all trajectory curves of points in  $B$ :

$$\begin{aligned} \text{tube}(S(F, t, B)) &= \{\mathbf{f}(\alpha) + M(\alpha)\mathbf{b} : \mathbf{b} \in B, 0 \leq \alpha \leq t\} \\ &= \cup \{F(\alpha)B : \alpha \in [0, t]\}. \end{aligned}$$

This coincides with the tubelike space region that accommodates  $B$ 's passage.

If sweptspace  $S(F, t, B)$  coincides with  $\text{tube}(S(F, t, B))$ , we call both  $S$  and the sweptspace *tubular*. Points in such a sweptspace clearly originate in  $B$ , in the sense that the tube is the union over all  $\mathbf{b}$ -trajectories,  $\mathbf{b} \in B$ .

The *sweep curve* is the translation curve,  $\mathbf{f}$ ; it is identifiable as the trajectory, or tube, of  $\{\mathbf{0}\}$ —i.e., as the image set,  $\mathbf{f}([0, t])$ , of  $[0, t]$ . A related subset of the sweptspace is the set  $S(t, \{\mathbf{0}\})$ , or, by understanding,  $S(t, \mathbf{0})$ ; we call this set the *swell*. It is the swept image of the singleton set,  $\{\mathbf{0}\}$ , and for a tubular sweep it agrees with the sweep curve, but for cumulative sweeps (below) it is more extensive.

$S$  is a *translational sweep* if  $M(t) = M(0)$  = the identity matrix; a *rotational sweep* if  $\mathbf{f}(t) = \mathbf{f}(0)$ ; and a *mixed sweep* if both  $M$  and  $\mathbf{f}$  vary. If  $S$  is translational and tubular, then its sweptspace is the Minkowski sum,  $S(t, B) = \mathbf{f}([0, t]) + B \equiv S(t, \mathbf{0}) + B \equiv \text{swell} + B$ ; in fact, this decomposition applies to all translational sweeps, even nontubular ones (below).

We describe sweep  $S$  as *cumulative*, and say it has a *memory*, if the sweptspace at any time  $t$  is more extensive than the corresponding tube. This is made exact by membership rules to be developed; informally, it means that points may enter a sweptspace at time  $t$  not only along point trajectories that have originated in  $B$  but also along point trajectories that have originated in any earlier sweptspace. For example, a point  $\mathbf{x}$  belonging to  $S(F, t_1, B)$  may be permitted to spawn member  $\mathbf{x}' = F(t_2)F(t_1)^{-1}\mathbf{x}$  of a later sweptspace,  $S(F, t_2, B)$ . It is possible to place time limits in such origination rules in a manner that gives exact sense to the notion of a memory of specified duration; we return to this matter in our final section.

<sup>2</sup> J. Meyer and M. Lavin, "Swept Volume of Polyhedral Models" (unpublished manuscript), IBM Research Division, Yorktown Heights, NY.

### Sweep types in OYSTER

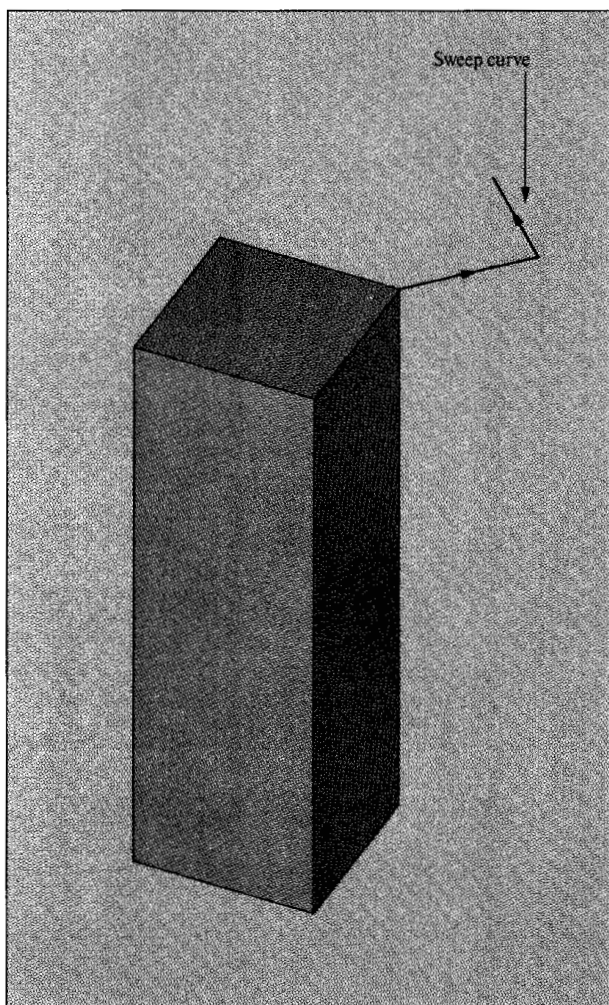
The preceding indicates that a very broad study of sweeps entails consideration of rotational, translational, and mixed sweep types, and of sweptspaces which are either tubular or cumulative and which, if cumulative, may have varying durations of memory. For the modeling needs of OYSTER, the focus is narrower.

The GDP modeler has an algorithm that creates a polyhedral approximation, under tubular mixed sweep, for the tube of a polyhedral body; by restriction, this algorithm makes available a precise representation for the polyhedral sweptspace that results from the tubular translational sweep of a polyhedron along a single line-segment sweep curve; and by iterating the latter process along finitely many line segments of a piecewise-linear sweep curve, taking the

output sweptspace from each stage (line segment) as the input body to the next, we achieve the special CTS used in OYSTER simulations.

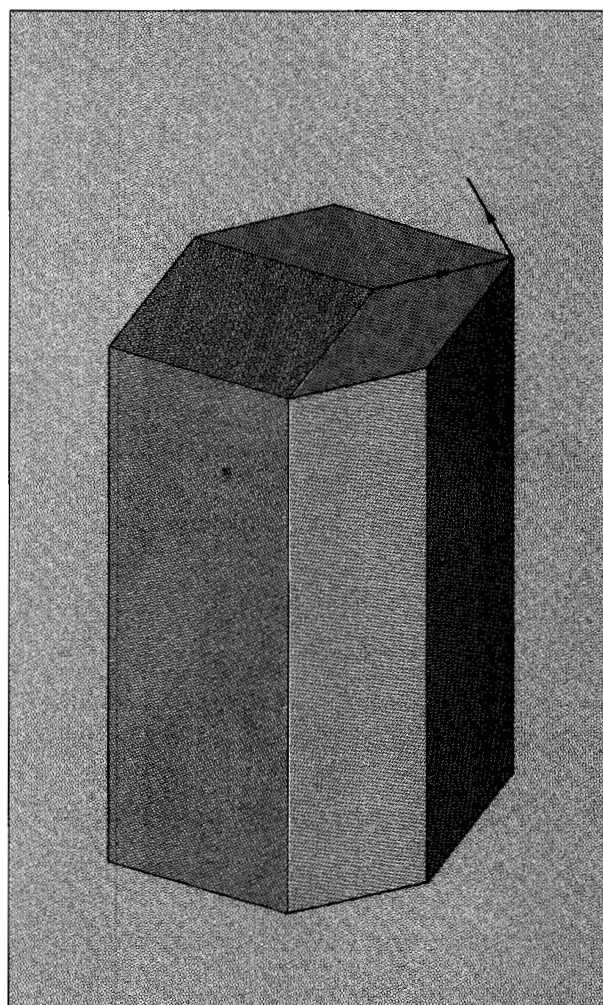
An example is given by the **Figure 1** sequence, which shows the generation of an octagonal prism from a square prism. In Figure 1(a), the sweep curve is seen as a simple two-segment curve in proximity to the input prism; the first curve segment prescribes a tubular sweep that develops the sweptspace of Figure 1(b), shown extending the original prism; sweeping this result along the second segment then determines the output prism, Figure 1(c), by inclusion of a new extension. In such a manner, by iterating the tubular translational sweep (TTS), one achieves a single CTS.

Figure 1(d) indicates another way of conceptualizing the result. It shows the *rays* of the sweep, i.e., the separate sweep



**Figure 1(a)**

Square prism from which an octagonal prism is to be generated by means of a two-segment sweep curve.



**Figure 1(b)**

Result of sweeping the square prism along the first segment of sweep curve.

curve segments treated (and stored by OYSTER) as vectors, and the swell—in this case a shaded parallelogram identifiable with the effect of the sweep upon a single point. The figure suggests how the output prism derives from the Minkowski sum of the swell and input prism—informally, by sliding the swell around the boundary of the input prism while maintaining the swell's orientation and (with respect to the swell) the point of coincidence.

To clearly understand the CTS as a Minkowski sum, consider the rays of the sweep to be given by vectors  $\mathbf{r}_1, \dots, \mathbf{r}_n$  ( $1 \leq n \leq N$ ); that is, suppose the time parameter values associated with endpoints of the linear segments of  $\mathbf{f}$  are  $0 = t_0 \leq t_1 \leq \dots \leq t_N$ , and define  $\mathbf{r}_n \equiv \mathbf{f}(t_n) - \mathbf{f}(t_{n-1})$ . Then, retaining the distinctions based on index, call set  $R = \{\mathbf{r}_1, \dots, \mathbf{r}_N\}$ , the *rayset* of the sweep. The first ray then prescribes a sweep that creates sweptspace

$$\begin{aligned} S(t_1, B) &\equiv S_{|\mathbf{r}_1|}(B) = \{\alpha \mathbf{r}_1 + \mathbf{b} : 0 \leq \alpha \leq 1, \mathbf{b} \in B\} \\ &= S_{|\mathbf{r}_1|}(\mathbf{0}) + B. \end{aligned}$$

And, if  $\mathbf{f}$  has more than one linear segment, a second ray iterates the process, developing

$$S(t_2, B) \equiv S_{|\mathbf{r}_1, \mathbf{r}_2|}(B) = S_{|\mathbf{r}_2|}(\mathbf{0}) + S_{|\mathbf{r}_1|}(\mathbf{0}) + B.$$

Continuing in this manner, it is apparent that the final CTS sweptspace is

$$\begin{aligned} S(t_N, B) &\equiv S_{|\mathbf{r}_1, \dots, \mathbf{r}_N|}(B) \equiv S_R(B) \\ &= S_{|\mathbf{r}_1|}(\mathbf{0}) + \dots + S_{|\mathbf{r}_N|}(\mathbf{0}) + B \equiv S_R(\mathbf{0}) + B. \end{aligned}$$

Observe of this CTS that swell  $S(t_N, \mathbf{0})$  has been given a new denotation,  $S_R(\mathbf{0})$ ; we abbreviate this to  $S_R$ ; and if  $R$  has only one ray,  $\mathbf{r}$ , we abbreviate further still, to  $\tilde{\mathbf{r}}$ , since the swell is then the line segment given by  $\tilde{\mathbf{r}} \equiv \{\alpha \mathbf{r} : 0 \leq \alpha \leq 1\}$ .

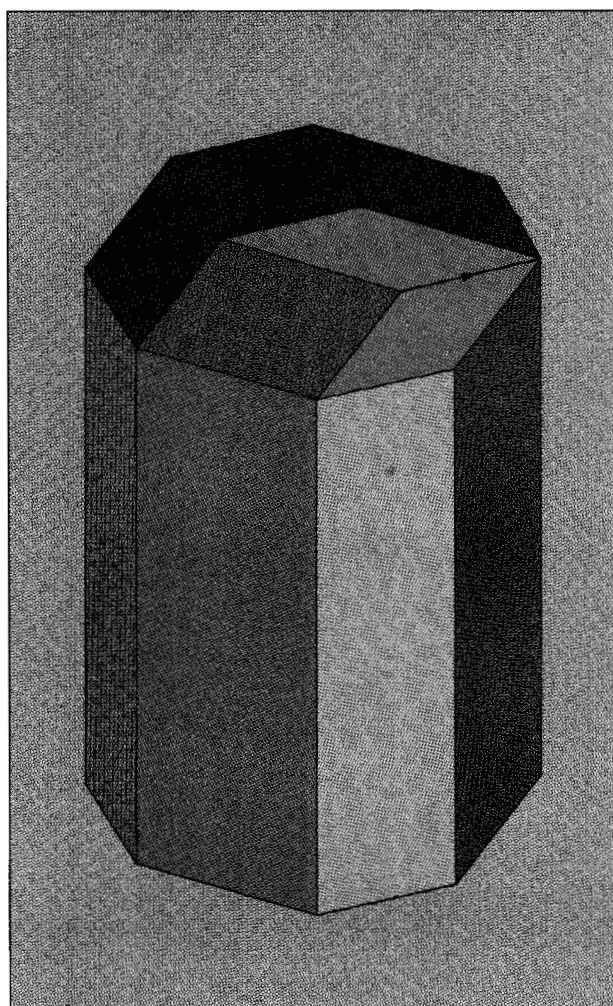


Figure 1(c)

Octagonal prism resulting from subsequent sweep along the second segment of sweep curve.

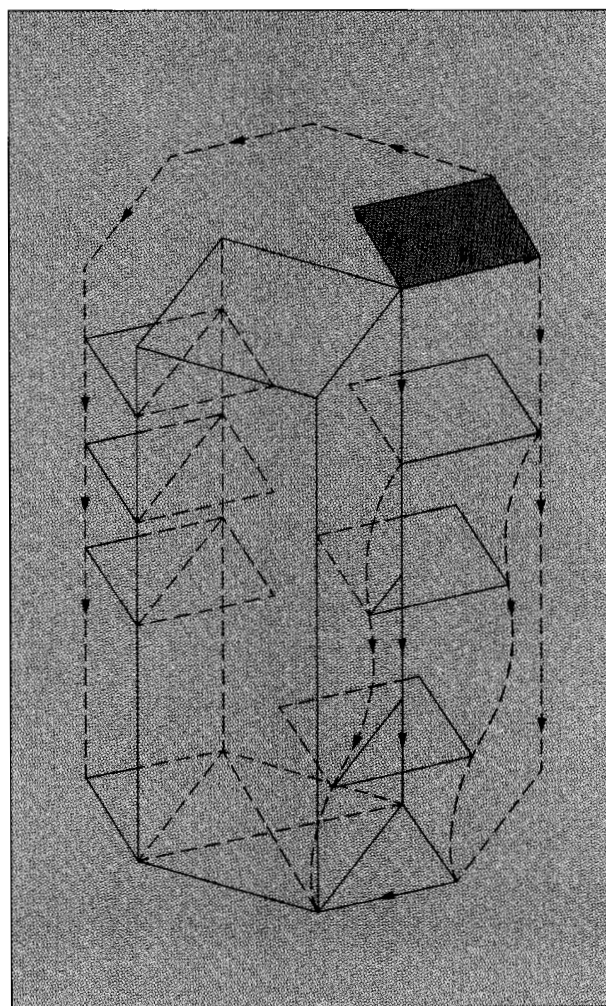


Figure 1(d)

Conceptualizing the octagonal prism as a Minkowski sum of the square prism with a parallelogram swell.

As indicated, one may express sweptspace  $S(t_N, B)$  as a sum which separates the contributions of the swell from those of the body—as  $S_R(B) = S_R + B$ . This extends an observation made earlier about the TTS to one that applies to a ray-based CTS; it is true of *every* translational sweep,  $S$ , that  $S(t, B) = S(t, \mathbf{0}) + B$ , and one may study translational shaping processes—disregarding the bodies to which they may be applied—by examining the shape of the swells that represent their effects.

Although swell  $S_R$  was introduced with regard to an order of rays reflecting the order of line segments along sweep curve  $\mathbf{f}$ , it is identified with the commutative set sum,  $S_R = \bar{r}_1 + \dots + \bar{r}_N$ , and thus is order-independent. It may be associated with *any* sweep curve derived from  $\mathbf{f}$  by permuting the order of its linear segments; when computing sweptspace  $S_R(B)$ , we are free to choose the permutation of rays that will minimize cost.

The preceding exhibits swells as Minkowski sums involving finitely many line segments. Such sets are known in mathematics as *zonotopes* [13–16]; limits of zonotope sequences are called *zonoids*—they appear in some surprisingly different contexts [19].

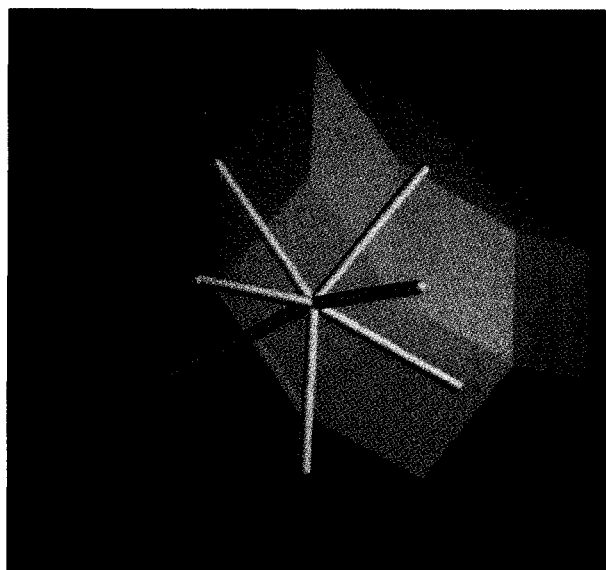
A *polytope* is the convex hull of a finite set of points. An alternative definition of a zonotope that de-emphasizes its line segment basis is that it is a centrally symmetric polytope having centrally symmetric facets of every order, where a set is *centrally symmetric*, or *centered*, if it reflects through one of its own points onto itself, in this exact sense:  $X$  is centered at  $\mathbf{c} \in X$  if  $\mathbf{p} \in X$  implies segment  $\overline{pp'} \subset X$ , where  $\mathbf{p}' \equiv 2\mathbf{c} - \mathbf{p}$  is the reflection of  $\mathbf{p}$  through  $\mathbf{c}$ . In other words—and this is a characterization we later employ— $X$  is centered at  $\mathbf{c}$  iff for every  $\mathbf{p} \in X$  and  $\alpha \in [0, 2]$ ,  $\mathbf{p} + \alpha(\mathbf{c} - \mathbf{p}) \in X$ .

Starting with a rayset, one may produce the zonotope which is its associated swell by applying the rayset, as a CTS prescription, to the origin. To go in the reverse direction, from an initial zonotope to a rayset which generates it, one may determine rays in this manner: Call the zonotope's edges equivalent if they are parallel; let one ray (directed either way, by choice) represent each such equivalence class; the rayset so determined generates a translationally congruent zonotope (its location being influenced by the direction choices made—see discussion of ray reversal below).

### Swells as shaping tools

Notation  $S_R$  and terms such as “swell” call attention to the underlying rays and the growth dynamic that may be associated with zonotopes. We now return to that focus—natural to the study of sweeping as a growth process.

When point  $\mathbf{x}$  is subjected to the CTS determined by rayset  $R$ , it “swells” into zonotope  $\mathbf{x} + S_R$ , and thus into swell  $S_R$  if  $\mathbf{x}$  is  $\mathbf{0}$ . This image set may be of surprising beauty. An example is the 56-faced swell shown in **Figure 2**. Based upon eight rays, this shaper has been used in some OYSTER



**Figure 2**

Depicted in the foreground are eight radially directed rods, one hidden from view, which indicate rays that define a set of translational sweeps. If applied in cumulative sequence to any initial point—such that each successive sweep is used to enlarge the swept volume passed to it by its predecessor—they produce a multifaceted zonotope that is translationally congruent to the background figure.

applications. As instanced in this case, the number of faces developing from  $N (\geq 3)$  rays in 3D is  $N(N - 1)$ —provided no three of the rays are linearly dependent.

**Figure 3** illustrates CTS rounding of an  $L$ -shaped polygon by swells which approximate a circle, first roughly and then more accurately. A four-element planar rayset develops the octagonal swell shown in **Figure 3(a)**, and an eight-element one develops the sixteen-sided polygonal swell shown in **Figure 3(b)**. The results show that the  $L$  develops a cover,  $S_R(L)$ , that has new edges around each original convex vertex—a simple approximation to rounding. The figure also illustrates the relations between swells as shapers and the grown or shrunken bodies that CTS creates. Output sweptspaces are shown in relation to the swells, in a manner that clarifies the Minkowski summation involved; swells seem to have acted upon input polygon  $L$  by sliding around its boundary,  $\sigma L$ , and displacing it into a new position—determined with respect to the original by the shape of the swell.

That a sweptspace,  $S_R(L)$ , is describable in terms of activity along boundary  $\sigma L$  is evident from the observation that points of the interior,  $L^\circ$ , cannot be swept into new territory without first crossing the frontier; to decide what new territory is to be annexed, it suffices to follow the frontier-crossing points, i.e., the points of  $\sigma L$ . This observation is formally derivable as an argument parallel to



In Figure 3, inward offset lines delineate the inward extent sweep.

(dual) "motion" of the well, not the tube of the original referring to the "tube" we mean the tube associated with this zonotope-shaped cutting tool). In what follows, when researchers may think of  $\sigma L$  as a control path and  $S_R$  as a contains a prescription for sweeping  $S_R$  (numerical control contains a prescription for sweeping  $L$ , but from another,  $\sigma L$  and  $S_R$  as the body to be swept; from one viewpoint,  $R$  tubular translational sweep that takes  $\sigma L$  as its sweep curve noting that  $S_R(\sigma L)$  is the tube of another sweep, namely the We complete the discussion concerning  $L$ 's boundary by best less straightforward and at worst more costly.

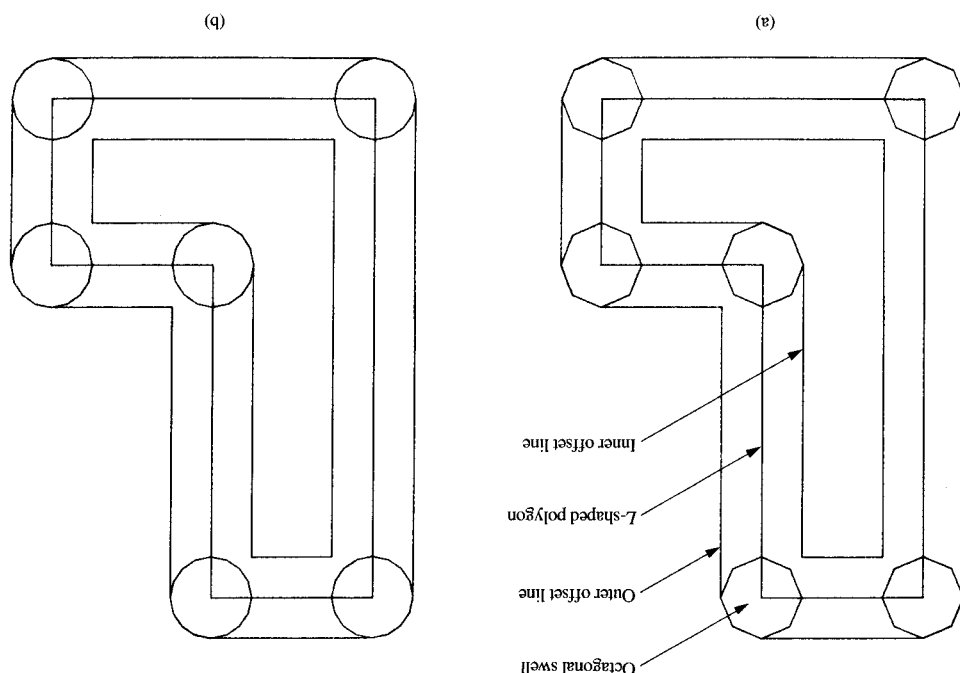
models employing space-filling representations would be at models that utilize boundary representations; their use in techniques are best suited for use in those geometric These boundary observations suggest that CTS shaping establishes that  $S_R(L) = L \cup S_R(\sigma L)$ .

"within  $d$  of" converts that argument into one which enlargement of nonempty set  $B$ , or set of points within  $d$  of  $L$   $E(d, B)$  denotes, for a space with metric  $\rho$ , the distance  $d$  one that shows that  $E(d, L) = L \cup E(d, \sigma L)$ —where

The simulation of rounding at the corners of  $L$  improves as the number of rays is increased. By taking sufficiently many sufficiently small rays, a planar well may be made to

The point of the well which always maintains tube has completely swallowed the well. boundary retains  $L$ 's angularity, a hint that on one side the across the tube from a faceted impression the opposite tube complement) the relics are left near  $L$ 's concave corners; passage; along the inward offset line (growth line for  $L$  image of itself in the outward line, a faceted relic of its at a convex corner (with respect to  $L$ ), it impresses a partial extent. As the well passes through the tube and takes a turn of the tube; outward offset lines delineate the outward

Figure 3 Planar rounding of an L-shaped polygon, showing swells and offset lines: (a) Rounding achieved by four rays (octagonal swell). (b) Rounding achieved by eight rays (16-sided swell).

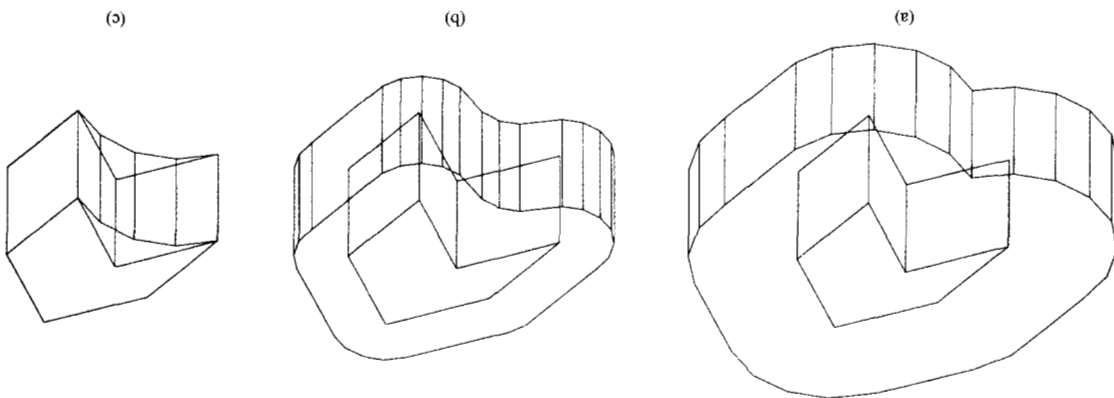


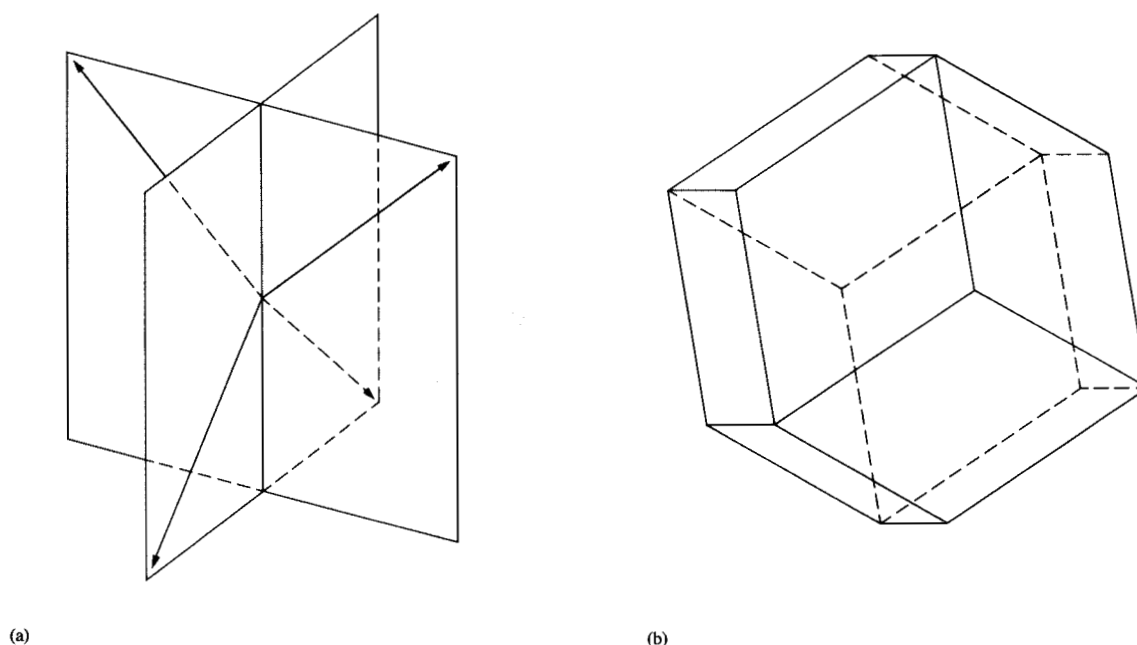
final thickness characteristics determined by rayset  $R$  and in a manner that rounds near *both* its concave and convex edges, one may grow  $B$  under control of rayset  $2R$  and shrink back under control of rayset  $R$ , i.e., may apply  $shr_R(gr_{2R}(*))$ —a combination with its own useful dual. These effects are illustrated in Figure 4, where a polygonal prism has been grown and then shrunk back under control of rays (in rayset  $R$ ) perpendicular to its axis. Figure 4(a) shows the sweepstakes after the first,  $gr_{2R}$ , stage. Here, the sharpness of original concavities has been propagated, but if inward "cusps" are then rounded by a  $shr_R$  operation, there results an intermediate sweepstakes, Figure 4(b), which is "halfway back" and "rounded everywhere"; shrinking "all the way back"—either from the Figure 4(a) sweepstakes using  $shr_R$  or from the Figure 4(b) sweepstakes using  $shr_R$ —creates the fill of Figure 4(c). A different aim might be to round along all vertical edges of the prism while leaving the central regions of each of its vertical faces in fixed locations; the sequence needed would be a simple extension of the above. We illustrate rounding under control of two 3D swells, the four-rayed rhombic dodecahedron of Figure 5 and the eight-rayed swell of Figure 2. When the latter is combined with an  $L$ -shaped cylinder, the CTS reshapes the cylinder as shown in Figure 6. And the effect of the four-rayed sweep, applied to a 3 by 3 by 3 lattice arrangement of unit cubes, is shown in Figure 7. The input lattice, Figure 7(a), was grown from a single point by a sequence involving three orthogonal rays, one application of  $gr$  and six applications of  $gr_i$ ; the reshaped lattice, Figure 7(b), restates in each of its 27 repeated "modules" what the four-ray sequence does to a single cube—or, with a slight twist, what the particular seven-ray sequence would do to a point.

•  $Grow\ L = gr_R(L) = S_R(L)$ .  
 •  $Shrink\ L = shr_R(L) = complement(gr_R(complement\ L))$ , achieving the Boolean complement of the grown Boolean complement of  $L$ .  
 •  $Layer\ L = lay_R(L) = gr_R(L) - L$ , where the minus sign denotes Boolean difference. We also call this result deposit, or the incremental growth,  $gr_i(L)$ .  
 •  $Erase\ L = er_R(L) = L - shr(L)$ . We also call this result the withdrawal, or the incremental shrinkage,  $shr_i(L)$ . Sweeps remove material layers as well as depositing them, since Boolean complementation introduces a pervasive duality. What to body  $B$  is growth is to its complement shrinkage, and what to body  $B$  is convex is to its complement concave. Sequences which combine growth and complementation under varying rayset control serve many special needs; we indicate two such:  $shr_R(gr_R(*))$  achieves fills along any edges where two faces meet concavely [by duality, the  $gr_R(shr_R(*))$  combination "sands off" along "sharp" edges where faces meet convexly]; and, in order to layer  $B$  with

approximate arbitrarily well a circular disk. Thus, in the  $xy$  plane, if  $N$  rays are defined by letting  $r_n$  be the ray of length  $l_n$  at angle  $(2n - 1)\pi/2N$  (counterclockwise from the  $+x$  axis), where  $l_n = 2 \sin^{-1}(\pi/2N)$ , then the swell is a  $2N$ -gon inscribing the unit circle and inscribed by a circle of radius  $\cos(\pi/2N)$ . The aspect ratio for the  $N$ th such swell, or ratio of its maximum to minimum diameter, is  $1/\cos(\pi/2N) \rightarrow 1$  (with value 1.04, for example, when  $N = 6$ ); this proves convergence to a disk. Regions corresponding to the following operational definitions may also be identified in Figure 3:

Figure 4 Growing and shrinking a polygonal prism: (a) Growth by a relative amount taken as 2—rounds convex corners. (b) Shrinking back the rest of the way—achieves a fill. (c) Shrinking back by a relative amount of 1—rounds concave corners.





**Figure 5**

Four-rayed swell used for single faceting: (a) Orientation of the four rays. (b) The generated swell, a rhombic dodecahedron, creates one new facet around edges directed in the  $x$ ,  $y$ , or  $z$  directions.

### Properties of ray-based swells

As a zonotope, swell  $S_R$  is known to be convex, centered, and compact (for Euclidean space, read closed and bounded)—say a *3C set*. We suggest here how these properties may be established as consequences of attribute preservation under set summation. Readers familiar with these arguments may wish to skip ahead to the discussion on the shape of planar swells.

#### $S_R$ is convex

Line segments are convex, so it suffices to show that set summing preserves convexity, i.e., that if  $A$  and  $B$  are convex, so is  $A + B$ . Let  $\mathbf{x}_1 = \mathbf{a}_1 + \mathbf{b}_1$  and  $\mathbf{x}_2 = \mathbf{a}_2 + \mathbf{b}_2$  be any two points of  $A + B$ , and consider an arbitrary convex combination of them, say  $\mathbf{x}_3 \equiv \alpha \mathbf{x}_1 + (1 - \alpha) \mathbf{x}_2$ . Then  $\mathbf{x}_3$  is clearly re-expressible as  $[\alpha \mathbf{a}_1 + (1 - \alpha) \mathbf{a}_2] + [\alpha \mathbf{b}_1 + (1 - \alpha) \mathbf{b}_2] \equiv \mathbf{a}_3 + \mathbf{b}_3$ . As a convex combination of  $A$ -points,  $\mathbf{a}_3$  is in  $A$ , and similarly, as a convex combination of  $B$ -points,  $\mathbf{b}_3$  is in  $B$ ; hence  $\mathbf{x}_3 \equiv \mathbf{a}_3 + \mathbf{b}_3$  is in  $A + B$ .

#### $S_R$ is centered, at $(\sum \mathbf{r}_n)/2$

Observe first the effect of set-summing two centered sets: If set  $A$  is centered at  $\mathbf{c}_a$ , and  $B$  is centered at  $\mathbf{c}_b$ , then  $A + B$  is

centered at  $\mathbf{c}_a + \mathbf{c}_b$ . For if  $\mathbf{a} + \mathbf{b}$  is a representative point of  $A + B$ , with  $\mathbf{a} \in A$ , and  $\mathbf{b} \in B$ , then  $\mathbf{c}_a + \mathbf{c}_b$  proves to be a center, because for any  $\alpha \in [0, 2]$ , point  $(\mathbf{a} + \mathbf{b}) + \alpha((\mathbf{c}_a + \mathbf{c}_b) - (\mathbf{a} + \mathbf{b})) \equiv [\mathbf{a} + \alpha(\mathbf{c}_a - \mathbf{a})] + [\mathbf{b} + \alpha(\mathbf{c}_b - \mathbf{b})]$  is expressed on the right side of this identity as a member of  $A + B$ .

Then, since line segment  $\tilde{r}$  is clearly centered at  $\mathbf{r}/2$ , it follows that  $S_R = \tilde{r}_1 + \dots + \tilde{r}_N$  is centered at

$$(1/2) \sum_{n=1}^N \mathbf{r}_n.$$

We note as a corollary that  $S_R$  is centered at the origin iff its rays sum to  $\mathbf{0}$ ; this is of interest when considering the distribution of new material that a body acquires during cumulative sweep.

#### $S_R$ is closed

More strongly, we show that if  $B$  is closed, so is  $S_R(B)$ ; the particular result comes by taking  $B = \{\mathbf{0}\}$ . We assume, since the argument is finitely repeatable, that  $R$  is the singleton,  $\{\mathbf{r}\}$ ; then  $S_R$  is  $\tilde{r} \equiv \{\alpha \mathbf{r} : 0 \leq \alpha \leq 1\}$ , and we show that set  $B^+ \equiv B + \tilde{r}$  is closed, i.e., that if  $\mathbf{x} \notin B^+$ , then  $\mathbf{x}$  is not a limit point of  $B^+$ : Take  $\mathbf{x} \notin B^+$  and consider line segment  $\text{seg} \equiv \overline{\mathbf{p}\mathbf{x}}$ , where  $\mathbf{p} = \mathbf{x} - \mathbf{r}$ ; it clearly avoids  $B$ , and since  $\text{seg}$



is a compact set outside the closed set  $B$ , there is an  $\epsilon > 0$  such that set  $\overline{seg} \equiv E(\epsilon, seg)$  avoids  $B$ . Now,  $\overline{seg}$  is a sausage-shaped set that covers on one end an  $\epsilon$ -neighborhood of  $\mathbf{p}$ ,  $E(\epsilon, \mathbf{p})$ , and on the other an  $\epsilon$ -neighborhood of  $\mathbf{x}$ ,  $E(\epsilon, \mathbf{x}) \equiv \mathbf{r} + E(\epsilon, \mathbf{p})$ . So  $\overline{seg}$  is of sufficient size that every line segment parallel to  $\mathbf{r}$  which spans its interior has length  $> |\mathbf{r}|$ ; it follows, if  $\mathbf{b} \in B$ , that  $\mathbf{b} + k\mathbf{r}$  can lie in  $E(\epsilon, \mathbf{x})$  for no positive  $k \leq 1$ , and thus that  $\mathbf{x}$  is not a limit point of  $B^+$ .

#### $S_R$ is bounded

Since  $\mathbf{x} \in S_R$  is expressible as a sum of rays with coefficients from  $[0, 1]$ ,  $|\mathbf{x}|$  is bounded by  $\sum |\mathbf{r}_n|$ , summed over all rays in  $R$ .

A tighter directional bound is available. For a convex body  $K$  that contains  $\mathbf{0}$ , the normalized support function of  $K$ , in direction  $\mu$ , is defined by  $h(K, \mu) \equiv \max \langle \mathbf{x}, \mu \rangle$ ,  $\mathbf{x} \in K$ ; this value represents the directional extent of  $K$  in the  $\mu$  direction, i.e., the distance between  $\mathbf{0}$  and that support plane of  $K$  which has its outward normal  $\mu$  pointing into the non- $K$  halfspace. When referring to zonotope  $S_R$ , in order to emphasize the relationship to rayset  $R$ , we denote this support function by  $H_R(\mu)$  and term it the *height* of the swell in direction  $\mu$ . If the swell is centered at  $\mathbf{0}$ , this is one half the convex set width in the  $\mu$  direction; in general, the width of the swell in the  $\mu$  direction is  $H_R(\mu) + H_R(-\mu)$ .

$S_R$  attains its height,  $H_R(\mu)$ , at the point which is the sum of the rays, if any, that point into the  $\mu$  halfspace—or at  $\mathbf{0}$  if there are none. For if  $R$  has no ray pointing into the  $\mu$  halfspace, then clearly  $\langle \mathbf{x}, \mu \rangle$  is maximized,  $\mathbf{x} \in S_R$ , by

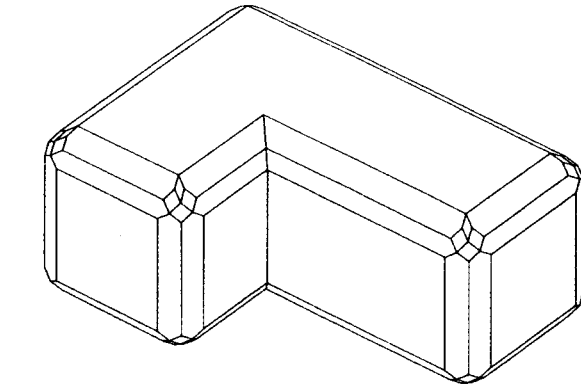
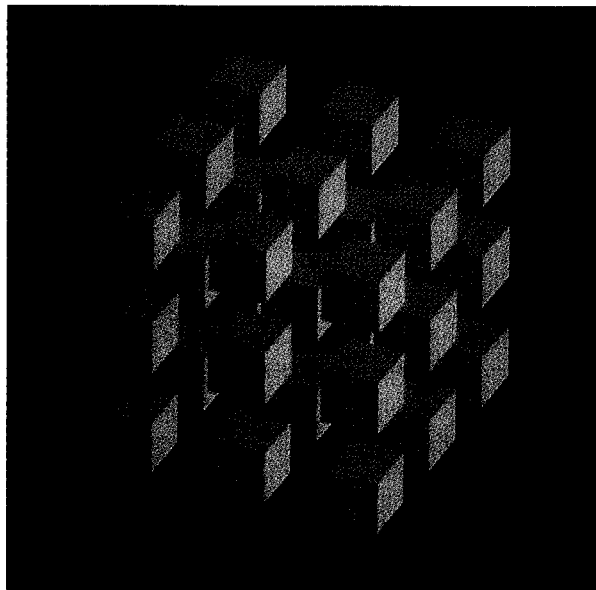


Figure 6

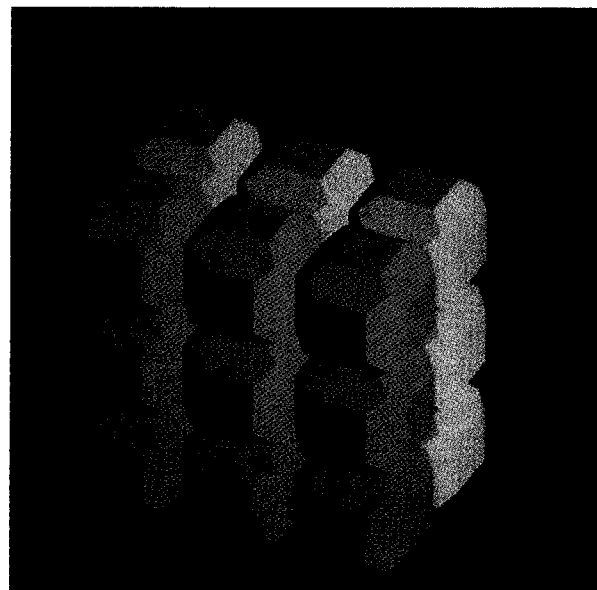
Polyhedron grown from an L-shaped prism by an eight-rayed CTS.

$\mathbf{x} = \mathbf{0}$ ; otherwise the maximum is sought with respect to representative points of form  $\mathbf{x} = \sum \alpha_n \mathbf{r}_n$ , with coefficients  $\alpha_n \in [0, 1]$ , and for such a point,  $\langle \mathbf{x}, \mu \rangle \leq \langle \mathbf{y}, \mu \rangle$ , where  $\mathbf{y} \equiv \sum \mathbf{r}_i$  summed over exactly those rays for which  $\langle \mathbf{r}_i, \mu \rangle > 0$ .

We also remark that  $H_R$  completely determines  $S_R$ .



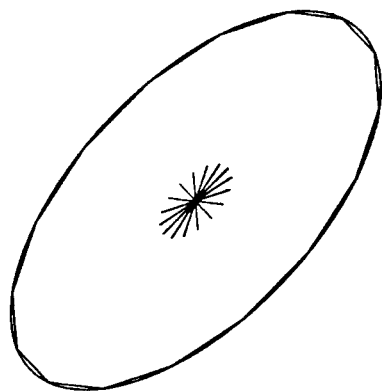
(a)



(b)

Figure 7

Dodecahedral shaping of a lattice of cubes: (a) Input lattice. (b) Shaped output.



**Figure 8**

Swell approximating a sheared ellipse (outermost boundary), showing rays (center) and approximating polygonal swell (inscribing the ellipse).

#### *The shape of planar swells*

We later observe a simple result—that reversing a ray’s direction translates a swell without altering its shape; also, since rays may be applied in any order, we may determine the shape of a swell by studying whatever sweep curve (among all those which differ only by the ordering or direction of their segments) most easily reveals it. In the planar case, there is always among such equivalent curves one whose discrete turnings are all in the same angular direction and of cumulative amount at most  $180^\circ$ . When we have chosen such a representative curve, the central symmetry property allows us to invert it through the center of symmetry and so exhibit as subset of the swell a simple closed curve formed of that representative and its inverted image; convexity then establishes the inclusion of the planar region inside the closed curve; and an argument based on swell heights,  $H_R(\mu)$ , establishes the exclusion of the planar region outside it, completing the swell’s characterization. As an example, any sweep curve which polygonally inscribes a semicircle generates a swell that inscribes the corresponding circle.

For a swell in a higher-dimensional space, the above technique permits determination of the shape of any of its planar projections by consideration of the projected planar raysets.

#### *Behavior of swells under affine transformations*

If  $L$  is a linear transformation and  $A$  is a subset of its domain, denote by  $LA$  the image set  $\{L(\mathbf{a}): \mathbf{a} \in A\}$ . Then the fact that everything is additively defined and  $L$  is linear results in this immediate conclusion:

$$LS_R = S_{LR}.$$

If sweptspaces are considered, the corresponding result is

$$LS_R(B) = S_{LR}(LB).$$

These observations have useful consequences:

- Approximations by swells. If  $R$  is a known rayset such that  $S_R$  approximates  $A$  (perhaps a sphere), then  $LR$  is a rayset such that  $S_{LR}$  approximates  $LA$  (perhaps an ellipsoid). As a 2D example, see **Figure 8**. A rayset comprising 16 rays originally equally spaced around a circle—and which, if taken as such, would have generated an approximately circular swell—has been transformed so that its swell approximates a sheared ellipse.

There are also swells to approximate the unit ball in any dimension [16]. The Hausdorff distance between sets  $X$  and  $Y$  is  $\rho(X, Y) \equiv \max\{\sup_{x \in X} \inf_{y \in Y} |x - y|, \sup_{y \in Y} \inf_{x \in X} |y - x|\}$ . Using this metric in a  $d$ -dimensional space, Betke and McMullen ([20], Section 4) bounded the infimum distance between the class of  $n$ -rayed zonotopes and the unit ball,  $B$ ; the inf lies between values  $\beta_d n^{-2}$  and  $\gamma_d n^{-2/(d-1)}$ , for constants  $\beta_d$  and  $\gamma_d$  independent of  $n$ . In 3D, the principal author tested spherical approximations based upon two experimental rayset sequences, one that employed rays pointing toward mesh points of an ever-refined geodesic dome, another whose rays pointed in ever-tighter spiral patterns around a “northern” hemisphere. The number of rays needed to achieve (statistically tested) aspect ratios less than 1.01 was determined—approximately 2500 under either sequence.

- Determination of the shape of a swell. If  $L$  projects into a plane, then the earlier remarks on determining the shape of planar swells apply to  $S_{LR}$ , permitting the full determination of the shape of this planar projection of  $S_R$ , and providing a very helpful aid to its full conceptualization. We have used such projections to design raysets that achieve particular growth profiles in each of three mutually orthogonal directions.
- Changing the shape of a swell to vary the thickness of deposited layers. The thickness of a deposited layer cumulatively grown under the prescription of rayset  $R$  is directionally dependent and given by the height function,  $H_R$ . If  $R$  were designed to induce unit thickening in each of three orthogonal directions, say the cardinal ones of  $xyz$  space, but different directional heights were preferred, such as 3, 45, and 100 (while preserving the faceting pattern characteristic of  $R$ ), they might be obtained by transforming the rays of  $R$  by application of a diagonal matrix having as diagonal entries those three values; use of this observation in the OYSTER application area often involves the same scalar at each diagonal entry, entailing straightforward scaling of each of the rays.

To scale the thickness by  $k$  in the arbitrary direction of unit column vector  $\mu$  (using  $\mu^T$  for its row transpose), and

to do so without altering thicknesses in directions orthogonal to  $\mu$ , one may apply transformation  $I - (1 - k)\mu\mu^T$ . To independently scale the thicknesses in three orthogonal directions given by the columns of orthonormal matrix  $U$ —and by three amounts given along the main diagonal of diagonal matrix  $D$ —one may apply  $UDU^T$ .

It remains to consider translations. The basic observation is that translating a swept body is equivalent to sweeping a translated body;  $\mathbf{v} + S_R(B) \equiv S_R(\mathbf{v} + B)$ , or, in expanded form,  $\mathbf{v} + (S_R + B) \equiv S_R + (\mathbf{v} + B)$ .

Ordering the latter sum as  $(\mathbf{v} + S_R) + B$  suggests the interpretation that body  $B$  has been acted upon by a translated swell, an effect that may be grasped in terms of shift of hangpoint, i.e., in terms of an origin shift that creates a new *effective* hangpoint. Thus, translating a swell by  $\mathbf{v}$  creates an effective hangpoint at  $-\mathbf{v}$ , indicating a distribution of deposited material corresponding to a shifted point of conceptual coincidence between the swell and the boundary track of the body.

If  $S_{R'}$  were to have the shape of  $S_R$  but an effective hangpoint at  $\mathbf{v}$  rather than  $\mathbf{0}$ , its layering effect would be given by  $\text{lay}_{R'}(B) \equiv \text{gr}_{R'}(B) - B = \text{gr}_R(B - \mathbf{v}) - B$ , and as such could be readily achieved.

An effective hangpoint at the swell's center of symmetry is achieved by translating by  $\mathbf{v} = (-\Sigma\mathbf{r})/2$ . One at swell boundary point  $\mathbf{w}$  is achieved by translating by  $-\mathbf{w}$ ; it is sometimes useful to choose such a boundary point with respect to a given direction,  $\mu$ , as a point at which height  $H_R(\mu)$  is attained.

Figure 9 shows the effects described. A layer has been grown from a rectangular slab, Figure 9(a), in two different ways that use the same swell but treat it in terms of a different hangpoint. When the effective hangpoint is taken at the swell's center of symmetry, we consider that it determines a *balanced* swell; it induces growth layers that have a balanced appearance, as in Figure 9(b). When the effective hangpoint is taken somewhat "beneath" the center of symmetry, deposits occur somewhat "above" the body and we call the swell a *climber*; in Figure 9(c) the hangpoint has been placed at the lower boundary of the swell, raising the developed layer to the high point shown. When the effective hangpoint is taken above the center of symmetry, we call the swell a *digger*.

#### Reversal of ray direction

Reversal of ray direction effects a translation. For if  $R'$  is produced from  $R$  by the reversal of a single ray,  $\mathbf{r}$ , then the associated swells differ by translation vector  $\mathbf{r}$ , i.e.,  $\mathbf{r} + S_{R'} = S_R$ . In other words,  $\mathbf{r} + (\overline{-\mathbf{r}} + X) = (\overline{\mathbf{r}} + X)$ ; but this is obvious, since

$$\begin{aligned}\mathbf{r} + \overline{-\mathbf{r}} &= \{\mathbf{r} + \alpha(-\mathbf{r}) : 0 \leq \alpha \leq 1\} = \{(1 - \alpha)\mathbf{r} : \dots\} \\ &= \{\beta\mathbf{r} : 0 \leq \beta \leq 1\} = \overline{\mathbf{r}}.\end{aligned}$$

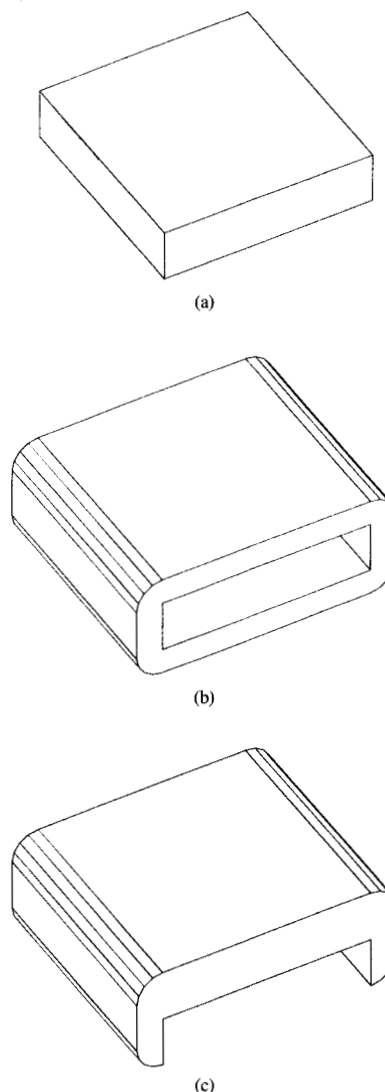


Figure 9

Varied layering dependent upon hangpoint location: (a) Underlying rectangular slab. (b) Balanced layer determined by a balanced swell (swell having hangpoint at the center of symmetry). (c) Rising layer determined by a "climber" (swell having hangpoint below the center of symmetry).

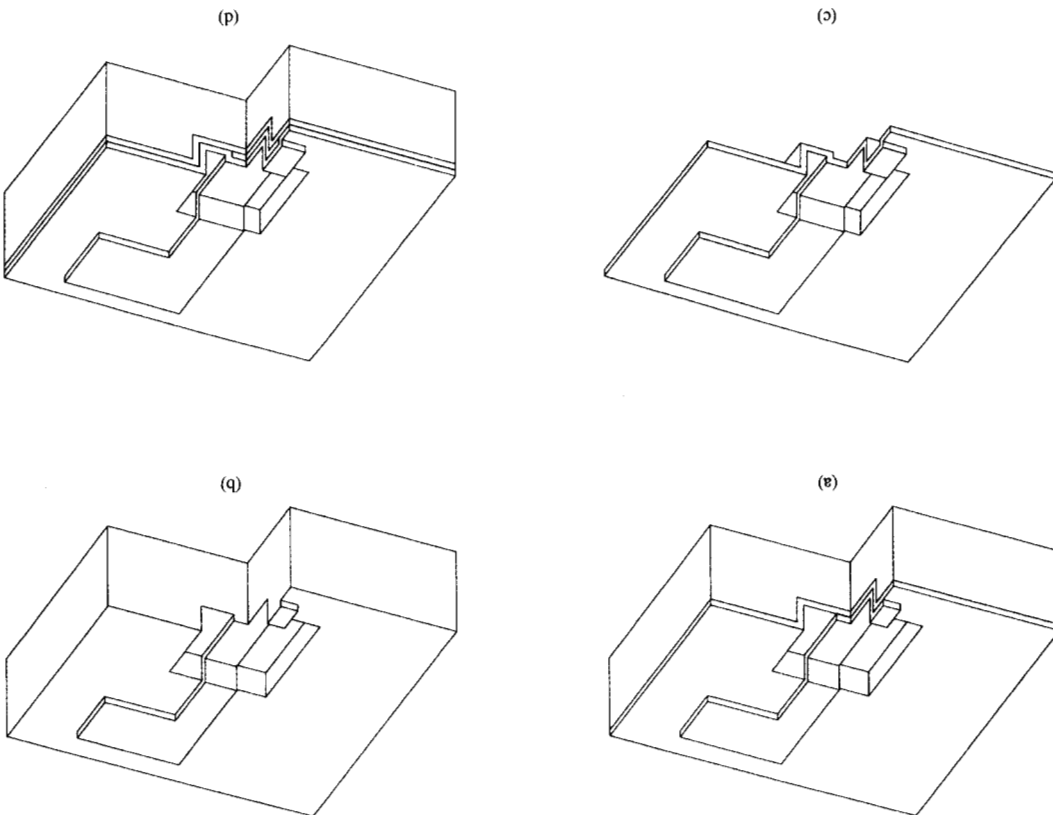
That ray reversal reduces to translation has important consequences:

- Shift of hangpoint. Since a translation is involved, the previous section shows that ray reversal effects a hangpoint shift. For example, imagine in  $xyz$  coordinates that

- Choice of cheaper sweep directions. If a polyhedron,  $P$ , multiple sweeps to compute it in the first place. it is computed—a much cheaper procedure than using be cheaply undone by translating the sweep space back after translation in the sweep space of swept objects, but that can replaced by a single ray; this replacement implies a in the same direction, the co-directed set then being opposite directions can be selectively reversed to all point defined by their sum. Then observe that multiple rays in application would have the same effect as the single ray store multiple rays in a common direction; their sequential Economy of ray storage. First note that it is wasteful to well with an interior hangpoint. reversing some but not all of  $R$ 's rays would produce a negative  $z$  halfspace and a well with hangpoint on top; rays then produces a digger,  $-R$ , with all of its rays in the has well with hangpoint (0) beneath it; reversing all of  $R$ 's climber  $R$  has all its rays in the positive  $z$  halfspace, hence

- Equivalence of curves. We observed earlier that the order of applying rays is irrelevant to the result, and we now see that the reversal of ray direction is irrelevant to the shape of the result (and just alters its position, by translation). We take sweep curves to be equivalent if they produce geometrically congruent swells (identical but for translation and rotation). Among the equivalent sweep curves are those which differ only in the order of their segments and the direction in which they run; in a sense, time reversal—doing last things first, in the opposite direction—does not alter the shapes produced by cumulative translational sweeps.
- Equivalence of curves. We observed earlier that the order of applying rays is irrelevant to the result, and we now see that the reversal of ray direction is irrelevant to the shape of the result (and just alters its position, by translation). We take sweep curves to be equivalent if they produce geometrically congruent swells (identical but for translation and rotation). Among the equivalent sweep curves are those which differ only in the order of their segments and the direction in which they run; in a sense, time reversal—doing last things first, in the opposite direction—does not alter the shapes produced by cumulative translational sweeps.

Figure 10  
Deposition steps performed by OYSTER: (a) Silicon device with several layers. (b) Composite union of device layers. (c) Layer grown by orthogonal three-element rayset. (d) Grown layer of Figure 10(c) shown in place conforming to previous layers.



$S_R$  as a convex hull

It is of interest to note that the vertices of a swell, and hence the swell itself (as convex hull of the vertex set), are expressible as a special sum over selected subsets of the rays.

We call  $X \subset R$  a *strong subset* of  $R$ , and  $v \equiv \sum_{x \in X} x$  a *strong sum*, if for some unit vector  $\mu$ ,  $\langle x, \mu \rangle > 0$  for every  $x \in X$ , and  $\langle r, \mu \rangle < 0$  for every  $r \in R - X$ .

Vector  $0$  is not a strong sum, because for every vector  $\mu$ ,  $0 = \langle 0, \mu \rangle$ .

Every strong sum is a vertex of  $S_R$ , and every vertex other than  $0$  is a strong sum. This follows from an argument, using the  $H_R(\mu)$  formulation, that shows a 1-to-1 correspondence between strong sums and support planes of  $S_R$  that contain exactly one (non- $0$ ) point. The above shows that zonotope  $S_R$  is the convex hull of  $\{0\} \cup \{\text{strong sums}\}$ .

We note, marginally, that  $0$  is a vertex of  $S_R$  iff  $R$  is a strong subset, and is a boundary point of  $S_R$  iff  $R$  is "almost strong," i.e., satisfies the strong conditions if the strict "greater than" ( $>$ ) is replaced by "greater than or equal to" ( $\geq$ ).

## OYSTER CTS applications

Recent developments in integrated circuit technology and predicted future trends point to the need for better modeling tools to help link a technology's design and manufacturing phases. The aim of fabrication modelers such as OYSTER is to model the geometric form of silicon devices at each stage of their manufacture, in order to aid device designers, who may derive operating characteristics from the models, and manufacturing engineers, who may use them to investigate allowable tolerance bounds.

OYSTER takes the designer's mask artwork and the manufacturing engineer's step-by-step description of the fabrication process as its inputs. It requires the same Boolean engine as more typical mechanical CAD/CAM applications, but must deal with more unusual shapes. In the silicon process world, the forms of materials that take shape on the wafer are continuous in layer-by-layer conforming patterns, suggestive of geological strata that flow and twist in unexpected and irregular ways. OYSTER attempts to capture the roundness of the layers and to accurately reflect their varying thicknesses, which depend upon the directional orientations of the various device elements.

A process step that occurs repeatedly is the deposition of a new layer of material upon an old. Although different methods of deposition and different material types effect distinctive layer shapes, all deposited layers do conform in shape to their underlying support. OYSTER achieves such new layers by applying the grow operation to a composite union of underlying layers and then Boolean differencing the result and underlying composite.

Figure 10 provides an example. Figure 10(a) shows a cutout view of an OYSTER model, representing a device which, during its manufacture, has developed several layers.

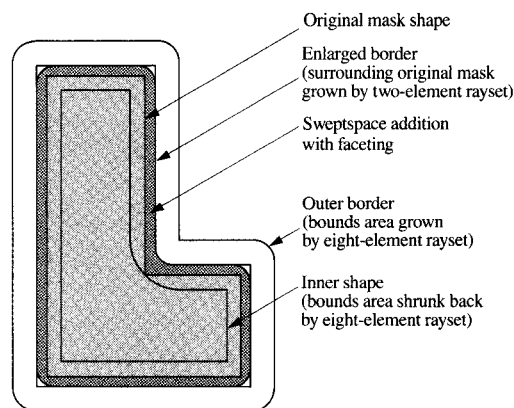


Figure 11

Mask shaping by repeated CTS operations.

The fabrication step to be illustrated is a uniform coating of the device by a new blanket material layer of specific thickness. To model the step, a union is created of all existing device components, as shown in Figure 10(b). The composite union might then be swept (in the simplest case) along three orthogonal directions and a new layer then be derived by a Boolean difference which removes the composite from the swept composite, Figure 10(c); the derived layer is shown in place, above the original layers, in Figure 10(d). Square edges and corners in the layer reflect the box-shaped swell employed. The combined use of Boolean and sweep operations illustrated in this example is typical of OYSTER algorithms which model fabrication steps.

In modeling silicon fabrication, it is important to take into account the shape rounding due to the lithography steps. The mask defining the L-shaped gate region layer used in Figure 10 can be considered a typical lithography mask shape as drawn by a device designer. Because of exposure tooling effects and material effects during the chemical development of the photoresist that has been exposed with the mask, the square corners become rounded, with a radius that is generally technology-dependent. The CTS sweep techniques permit us to facet around the corners and simultaneously grow or shrink the mask to compensate for fabrication effects. Figure 11 shows how an appropriate combination of grow and shrink operations with the rayset of Figure 5(a) is used to grow the original mask uniformly over the boundary while inducing faceting at both convex and concave corners.

Figure 12 shows the effect of combining planar mask faceting and deposition faceting to achieve more realistic

$S_R$  as a convex hull

It is of interest to note that the vertices of a swell, and hence the swell itself (as convex hull of the vertex set), are expressible as a special sum over selected subsets of the rays.

We call  $X \subset R$  a *strong subset* of  $R$ , and  $v \equiv \sum_{x \in X} x$  a *strong sum*, if for some unit vector  $\mu$ ,  $\langle x, \mu \rangle > 0$  for every  $x \in X$ , and  $\langle r, \mu \rangle < 0$  for every  $r \in R - X$ .

Vector  $0$  is not a strong sum, because for every vector  $\mu$ ,  $0 = \langle 0, \mu \rangle$ .

Every strong sum is a vertex of  $S_R$ , and every vertex other than  $0$  is a strong sum. This follows from an argument, using the  $H_R(\mu)$  formulation, that shows a 1-to-1 correspondence between strong sums and support planes of  $S_R$  that contain exactly one (non- $0$ ) point. The above shows that zonotope  $S_R$  is the convex hull of  $\{0\} \cup \{\text{strong sums}\}$ .

We note, marginally, that  $0$  is a vertex of  $S_R$  iff  $R$  is a strong subset, and is a boundary point of  $S_R$  iff  $R$  is "almost strong," i.e., satisfies the strong conditions if the strict "greater than" ( $>$ ) is replaced by "greater than or equal to" ( $\geq$ ).

## OYSTER CTS applications

Recent developments in integrated circuit technology and predicted future trends point to the need for better modeling tools to help link a technology's design and manufacturing phases. The aim of fabrication modelers such as OYSTER is to model the geometric form of silicon devices at each stage of their manufacture, in order to aid device designers, who may derive operating characteristics from the models, and manufacturing engineers, who may use them to investigate allowable tolerance bounds.

OYSTER takes the designer's mask artwork and the manufacturing engineer's step-by-step description of the fabrication process as its inputs. It requires the same Boolean engine as more typical mechanical CAD/CAM applications, but must deal with more unusual shapes. In the silicon process world, the forms of materials that take shape on the wafer are continuous in layer-by-layer conforming patterns, suggestive of geological strata that flow and twist in unexpected and irregular ways. OYSTER attempts to capture the roundness of the layers and to accurately reflect their varying thicknesses, which depend upon the directional orientations of the various device elements.

A process step that occurs repeatedly is the deposition of a new layer of material upon an old. Although different methods of deposition and different material types effect distinctive layer shapes, all deposited layers do conform in shape to their underlying support. OYSTER achieves such new layers by applying the grow operation to a composite union of underlying layers and then Boolean differencing the result and underlying composite.

Figure 10 provides an example. Figure 10(a) shows a cutout view of an OYSTER model, representing a device which, during its manufacture, has developed several layers.

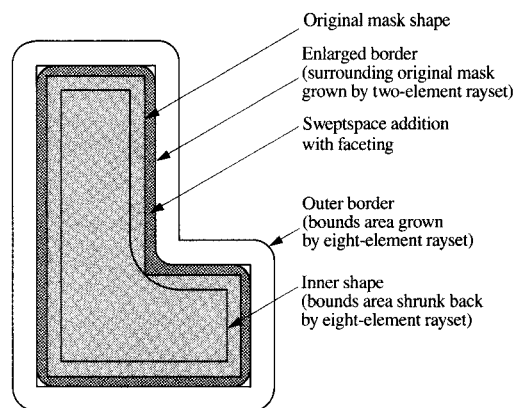


Figure 11

Mask shaping by repeated CTS operations.

The fabrication step to be illustrated is a uniform coating of the device by a new blanket material layer of specific thickness. To model the step, a union is created of all existing device components, as shown in Figure 10(b). The composite union might then be swept (in the simplest case) along three orthogonal directions and a new layer then be derived by a Boolean difference which removes the composite from the swept composite, Figure 10(c); the derived layer is shown in place, above the original layers, in Figure 10(d). Square edges and corners in the layer reflect the box-shaped swell employed. The combined use of Boolean and sweep operations illustrated in this example is typical of OYSTER algorithms which model fabrication steps.

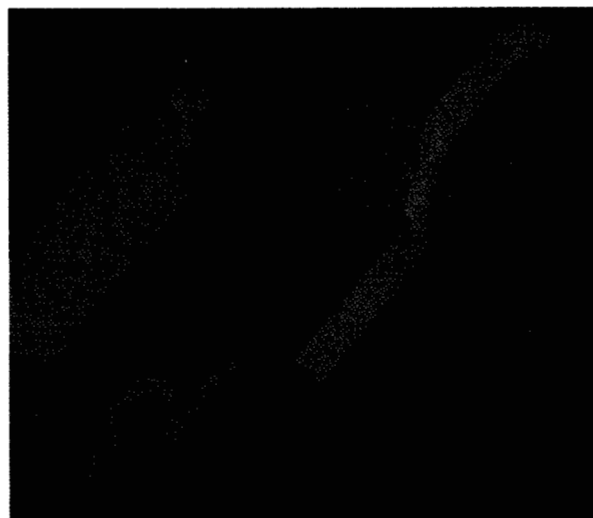
In modeling silicon fabrication, it is important to take into account the shape rounding due to the lithography steps. The mask defining the L-shaped gate region layer used in Figure 10 can be considered a typical lithography mask shape as drawn by a device designer. Because of exposure tooling effects and material effects during the chemical development of the photoresist that has been exposed with the mask, the square corners become rounded, with a radius that is generally technology-dependent. The CTS sweep techniques permit us to facet around the corners and simultaneously grow or shrink the mask to compensate for fabrication effects. Figure 11 shows how an appropriate combination of grow and shrink operations with the rayset of Figure 5(a) is used to grow the original mask uniformly over the boundary while inducing faceting at both convex and concave corners.

Figure 12 shows the effect of combining planar mask faceting and deposition faceting to achieve more realistic





(a)



(b)

**Figure 12**

Deposition steps performed by OYSTER with edge and mask faceting: (a) Deposited layers with one edge facet. (b) Deposited layers with two edge facets.

device shaping. Both masks that were used, to define the well area and the gate shape, have been rounded, as may be seen in Figure 11. A transmitted rounding is seen in corresponding faceted shapes in the device, as shown in Figure 12(a). Compare the shapes in Figure 12(a) to those of Figure 10(a). The well has been etched vertically downward in both figures, but the material layer beneath the gate has been deposited with one edge facet by using the rayset of Figure 5(a). The *L*-shaped gate region has been applied in a blanket layer and vertically etched as in Figure 10(a), but note that its shape, when dipping into the well, conforms to the edge faceting of the material layer beneath.

Figure 12(b) was created using the same rounded masks as in Figure 12(a), but the depositions were done with two edge facets instead of one, using a swell designed for double faceting, that of Figure 2. Details of rayset construction and swell hangpoints are made transparent to the OYSTER user, who merely selects the number of facets by setting a global parameter.

Further applications occur in other OYSTER fabrication step simulations. Thus, for etching, a very thin blanket layer is first grown and then intersected with the complement of the composite union; this identifies which portions of the device are exposed to the etchant. Similar methods, which create and use temporary intermediate layers, give OYSTER an ability to simulate growth of thermal oxides. Other applications of CTS are being considered for modeling effects particular to the type of deposition.

We chose to apply sweeping techniques for their many advantages. However, the method does raise issues of

complexity that we have yet to fully explore; we observe here only that CTS is a good candidate for parallel processing, since particular "forward components" of a solid may be swept in parallel and then subjected to a summarizing Boolean union.

### CTS for general curves

We have discussed applications of the ray-based CTS, which is based on sweeps along piecewise-linear curves. Families of such curves provide the limit sequences that establish the general class of rectifiable curves. Any particular such limit sequence, say of piecewise-linear curves  $C_i$  to establish the rectifiability of curve  $C$ , would carry with it a natural sequence of zonotopes,  $Z_i = S_{R_i}$ , say, and the  $Z_i$  sets would approach (under the Hausdorff metric) a limiting zonoid,  $Z$ , which is the obvious choice among sets to define as swell  $S_C$ . Below, we develop this idea in terms of our earlier notation and usage.

Consider a rectifiable curve in Euclidean  $N$ -space, parameterized over  $[0, 1]$  by a continuous function,  $C$ ; let  $\{\dots, (a_i, b_i), \dots\}$  be a finite sequence of disjoint  $[0, 1]$  subintervals and let  $R \equiv \{C(b_i) - C(a_i) : i = 1, \dots\}$  be its corresponding rayset. We say that  $C$  supports  $R$ , or that  $R$  is based on  $C$  (is  $C$ -based), and, if the association with  $C$  is understood, we say that  $R$  is based on  $\{\dots, (a_i, b_i), \dots\}$ . Then the swell of  $C$  (or  $C$ -swell) is defined by

$$S_C = \text{closure}(\cup \{S_R : C \text{ supports } R\}).$$

Under this definition, there are infinitely many ray-based swells contained in  $S_C$ —any parallelogram generated from

two supported rays, any parallelepiped generated from three, etc. Figure 13 illustrates a curve,  $C$ , and one such parallelepiped.

We remark, without proof, that  $S_C$  is the limit set of sequence  $S_{R_i}$ , having taken rayset  $R_i$  in association with the  $i$ th partition of  $[0, 1]$  in any refinement sequence that establishes  $C$ 's arclength.

*Swell-equivalence* is introduced to identify curves, under this choice of equivalence relations:  $S_1 \sim S_2$  if swells  $S_1$  and  $S_2$  differ at most by translation and rotation; and  $C_1 \sim C_2$  if  $S_{C_1} \sim S_{C_2}$ .

We say a collection of disjoint domain subintervals,  $\{\dots, (a_j, b_j), \dots\}$ , may be *refined* by splitting one or more of its members into several; it may be *completed*, extended to fully span the  $[0, 1]$  domain, by adjoining each missing subinterval. Thus  $\langle(0, 0.2), (0.3, 0.7)\rangle$  may be refined in many ways, perhaps to  $\langle(0, 0.1), (0.1, 0.2), (0.3, 0.7)\rangle$ , but may be completed in only one, as  $\langle(0, 0.2), (0.2, 0.3), (0.3, 0.7), (0.7, 1)\rangle$ .

If rayset  $R$  is  $C$ -based with respect to an incomplete sequence and is extended by adjoining the missing rays from its completed sequence, then the extended rayset, denoted by  $R^*$ , is called *complete*, or the  $R$  *completion*. If  $R'$  is based on a refinement of the domain subintervals associated with  $R$ , it is called an  $R$  *refinement*.

We make these observations:

- The new notation extends the old consistently; if  $C$  is any one of the curve segments obtainable by stringing together, start-to-end in any order, all directed line segments associated with the rays of  $R$ , then  $S_C = S_R$ .
- One curve equivalent to  $C$  is the translate,  $\tilde{C} \equiv C - C(0)$ , having  $\tilde{C}(0) = 0$ .
- Since each singleton,  $\{C(t) - C(0)\}$ , is clearly a  $C$ -based rayset,  $\tilde{C} \subset S_C$ .
- Operations on curve segments that preserve swell-equivalence include translation; rotation; repositioning, start-to-end, the members of a finite family of  $C$ -subsegments that span  $C$ ; first reversing the direction of some, and then repositioning the members of such a family. To briefly elaborate:

The affine behavior is this: for a linear transformation,  $L$ ,

$$LS_C = S_{LC},$$

and for a translation,

$$v + S_C = S_{(v+C)}.$$

Reversal of direction [assuming, without loss of generality, that  $C(0) = 0$  and  $\bar{C}(t) \equiv -C(t)$ ]:

$$S_{\bar{C}} = -S_C = -C(1) + S_C \equiv \bar{C}(1) + S_C.$$

Segmentation into subsegments  $C_1, \dots, C_n$ :

$$S_C = S_{C_1} + \dots + S_{C_n}.$$

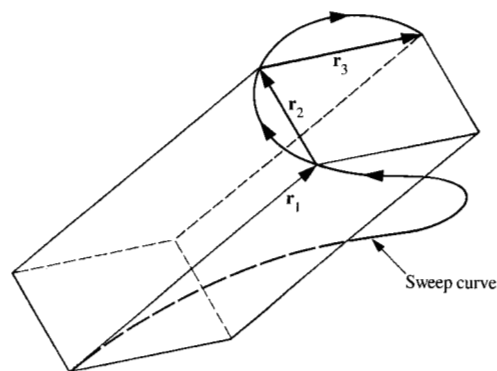


Figure 13

A general sweep curve and one of the infinitely many parallelepipeds contained by its swell.

- The nesting of completed refinements.

If  $R^*$  completes  $R$ , then, clearly,  $S_R \subset S_{R^*} \subset S_C$ ; and if  $R'$  refines  $R$ , then  $S_R \subset S_{R'} \subset S_C$ . The latter follows from the fact that, using the refinement structure, each ray  $r_i$  of  $R$  may be re-expressed as a sum,

$$\sum_{j=1}^{k_i} r_i^{j'}$$

over rays of  $R'$ , representative sums from  $S_R$  being similarly re-expressed as belonging to  $S_{R'}$ .

If  $R_1$  and  $R_2$  are  $C$ -based, they have a common completed refinement  $R$ ; it may be constructed in a unique way by completing that refinement of  $[0, 1]$  which involves exactly the division points associated with either  $R_1$  or  $R_2$ . Then, for  $i = 1$  or  $2$ ,  $S_{R_i} \subset S_R \subset S_C$ .

- $S_C$  is convex.

Let  $p$  and  $q$  be points of  $S_C$ , and  $\alpha \in [0, 1]$ . We must show that  $x \equiv \alpha p + (1 - \alpha)q \in S_C$ . Take  $p_i \rightarrow p$  and  $q_i \rightarrow q$ , with  $p_i$  and  $q_i$  in a common  $C$ -based rayset  $R_i$ . Since  $p_i$  and  $q_i$  belong to the convex zonotope,  $S_{R_i}$ , so does the point  $x_i \equiv \alpha p_i + (1 - \alpha)q_i$ ; and  $x_i \rightarrow x$ ; thus, since  $S_C$  is closed,  $x \in S_C$ .

- $S_C$  is centered, at  $c \equiv (C(1) - C(0))/2$ .

Note first that for any complete  $C$ -based  $R$ ,  $S_R$  is centered at  $(C(1) - C(0))/2$ , since it is centered at half the sum of its rays, which for a complete rayset telescopes to the stated value. Then, to complete the proof, we show that, if  $p \in S_C$  and  $\alpha \in [0, 2]$ , then  $p + \alpha(c - p) \in S_C$ . If  $p$  lies in some complete  $C$ -based  $R$ , then the observation above gives the result; otherwise, an argument strictly parallel to the above convexity argument shows that

$\mathbf{q} \equiv \mathbf{p} + \alpha(\mathbf{c} - \mathbf{p})$  is approached, through  $S_C$ , by  $\mathbf{q}_i \equiv \mathbf{p}_i + \alpha(\mathbf{c} - \mathbf{p}_i)$  and hence lies in  $S_C$ .

As corollary,  $S_C$  is centered at the origin iff  $C$  is a closed curve.

The above two properties (convexity and centrality at . . .) are helpful in understanding what regions  $S_C$  must include. Assume, for convenience, that  $C(0) = \mathbf{0}$ , and note that when  $\mathbf{p}$  is taken as  $C(t)$ , it reflects through the  $C(1)/2$  center into  $\mathbf{q} \equiv C(1) - C(t)$ , and as  $\mathbf{p}$  describes the  $C$  curve  $\mathbf{q}$  describes the reflected curve  $\tilde{C}$ . Then the swell of rayset  $R \equiv \{\mathbf{p}, \mathbf{q}\}$  is bounded by the parallelogram having vertices at  $\mathbf{0}$ ,  $\mathbf{p}$ ,  $\mathbf{p} + \mathbf{q} \equiv C(1)$ , and  $\mathbf{q}$ , and as  $t$  varies it moves through space "annexing territory for"  $S_C$ . (For planar curve  $C$  that turns in one direction through at most  $180^\circ$ , these observations confirm one made previously: that  $S_C$  consists of that planar region bounded by the closed curve formed from  $C$  and its central inversion.)

- $S_C$  is closed—by definition.

If closure were not part of the definition, then such a curve as the following would not have a closed swell (in fact its  $S_C$  would contain only two of its boundary points): In the  $xy$  plane, let the  $x$ -axis interval  $I_k \equiv [1 - 2^{-k}, 1 - 2^{-(k+1)}]$ ,  $k \geq 0$ , be the domain of  $C_k$ , a diameter  $2^{-(k+1)}$  semicircle (say in the  $y \geq 0$  halfplane if  $k$  is even, and in the  $y \leq 0$  halfplane if  $k$  is odd), and let  $C$  be  $\bigcup_{k=0}^{\infty} C_k$  together with point  $(1, 0)$ . Without closure, the swell of this snakelike curve would be the open disk of radius 0.5 centered at  $(0.5, 0)$  together with two boundary points,  $(0, 0)$  and  $(1, 0)$ ; closure brings in the entire boundary.

- $S_C$  is bounded by the arclength of  $C$ . If  $\mathbf{p} \in S_C$  is not  $\mathbf{0}$  and  $\mu$  is its unit vector, we have

$$|\mathbf{p}| \leq t_R(\mu) \equiv \sum_i \langle \mu, \mathbf{r}_i \rangle,$$

where the sum is taken over the rays in  $R$  that point into the  $\mu$  halfspace. Expressing this sum in terms of the  $C$  parameterization, as

$$\sum_i \langle \mu, C(\mathbf{b}_i) - C(\mathbf{a}_i) \rangle,$$

exhibits  $\mathbf{p}$  as bounded by  $C$ 's arclength, the least upper bound (lub) over all  $C$ -based finite sums

$$\sum_{n=1}^N |C(\mathbf{b}_n) - C(\mathbf{a}_n)|.$$

This arclength bound clearly applies to arbitrary points in  $S_C$ , as limit points of such  $\mathbf{p}$ 's.

The  $\mu$ -height of  $S_C$  may be defined, for a unit vector  $\mu$ , by  $H_C(\mu) \equiv \text{lub } H_{R_i}(\mu)$ , where the least upper bound is taken over all  $C$ -based  $R_i$ . This is the total variation of function

$$g(t) \equiv \max \{0, \langle \mu, C(t) \rangle\},$$

also expressible as

$$\int_0^1 \max \{0, \langle \mu, dC/dt \rangle\} dt.$$

- $S_C$  has been observed to be a 3C set, convex, centered, and compact. Are 3C sets always derivable as swells? No. One example of a three-dimensional 3C set which is not a swell is the regular octahedron: All 3D zonotopes have faces with evenly many sides, but the octahedron's faces are triangular. In 0-, 1-, or 2-space, however, one can find a swell to generate any 3C set:

- A zero-dimensional 3C set is a point, and is swell-equivalent to  $S_C$  for  $C(t) \equiv \mathbf{0}$ .
- A one-dimensional 3C set is a closed line segment, say of length  $k$ . It is swell-equivalent to  $S_C$  for  $C(t) = t(k\mu)$ , any unit vector  $\mu$ .
- A two-dimensional 3C set,  $A$ , has any semi-perimeter,  $C$ , as a generating curve; i.e.,  $A = S_C$ . This may be shown as follows:

Assume, without loss of generality, that  $C(t)$ ,  $t \in [0, 1]$ , parameterizes a semi-perimeter beginning at  $C(0) = \mathbf{0}$ .

To show  $A \subset S_C$ : If  $\mathbf{0} \neq \mathbf{x} \in A$  and  $\mathbf{y} \in (\sigma A) \cap$  (half-line extension of segment  $\bar{x}$ ), then for some  $\alpha$  and some  $t$ , both  $\in [0, 1]$ ,  $\mathbf{x} = \alpha\mathbf{y}$ , and either  $\mathbf{y} = C(t) \equiv \mathbf{p}$  or  $\mathbf{y} = C(1) - C(t) \equiv \mathbf{q}$ ; this establishes  $\mathbf{x} \in S_{\{\mathbf{p}, \mathbf{q}\}} \subset S_C$ .

Finally, to show complement( $A$ )  $\subset$  complement( $S_C$ ): If  $\mathbf{x} \notin A$ , then let  $\mu$  be its unit vector and let  $\mathbf{a}_x$  be its nearest point in  $A$  (which exists, because  $A$  is compact). Because of its convexity,  $A$  has a support plane,  $P$ , which passes through  $\mathbf{a}_x$ , is perpendicular to  $\mu$ , and locates  $\mathbf{x}$  in the open non- $A$  halfspace,  $\{\mathbf{p} : \langle \mu, \mathbf{p} \rangle > \langle \mu, \mathbf{a}_x \rangle\}$ .  $P$  may be assumed to support curve  $C$ , say at  $\mathbf{a}_x = C(t_0)$ ; if  $P$  supports  $C$ 's centrally reflected image, the argument might be rephrased in terms of the central reflection of  $\mathbf{x}$ . Then the directional variation of  $C$  in the  $\mu$  direction is realized by  $|\langle \mu, C(t_0) \rangle| = H_C(\mu)$ . Combining these observations, we have  $\langle \mu, \mathbf{x} \rangle > \langle \mu, \mathbf{a}_x \rangle = \text{lub} \{ \langle \mathbf{v}, \mu \rangle : \mathbf{v} \in S_C \}$ , hence  $\mathbf{x} \notin S_C$ .

The above completes a phase of discussion that has associated general curves with zonoids, just as an earlier phase associated piecewise-linear curves with zonotopes; the key to the association in each case has been the use of a new sweep type—the cumulative translational sweep—that directly links a curve to a "swell" (read zonoid).

We mention a further generalization before concluding: A limited-memory CTS, having memory of duration  $d$ , may be defined with clear intuitive sense in terms of a sweep curve,  $C(t)$ , that has been parameterized by arclength, say by  $t \in [0, T]$ : Let  $d$  be a continuous real-valued function on  $[0, T]$ , satisfying  $0 \leq d(t) \leq t$  [for the simplest case—memory of fixed duration— $d(t)$  might be  $\min\{D, t\}$ , for some constant,  $D$ ]. Let  $C_{[d,s]}$  be the segment of  $C$  defined over parameter subinterval  $t \in [s - d(s), s]$ . The CTS with memory of duration  $d$  that is defined under these circumstances is that sweep which generates as its swell

$$S_{[C,d]} \equiv \text{closure} (\cup \{S_{C_{[d,s]}} : 0 \leq s \leq T\}).$$

Swells of a limited-memory CTS lack the properties of convexity and central symmetry we have come to expect; other properties are also lost to the limited memory, such as the permutability of curve subsegments and the expressibility of the swell as a Minkowski sum of subordinate swells associated with a curve partitioning. In spite of these "failings," which make the limited-memory CTS more difficult to analyze than its unlimited correspondent, we believe the sweep type has considerable theoretical interest—that it may provide a conceptual tool for thinking about (even simulating) those phenomena in which temporal events may be associated with limited spatial effects. The richness of the constructs involved—permitting, for example, time-varying duration functions,  $d$ —suggests to us an area worthy of further study.

### Concluding remarks

We have described the CTS, illustrated its importance, and discussed many of its significant properties. Here we raise some open questions of both theoretical and practical importance.

Our work has shown how CTS curves determine shaping bodies that are zonoids, but—except for the zonotope subclass—it has not solved the equally important problem of how a general zonoid may determine an associated CTS curve. The solution must entail a plan, given a convergent zonotope sequence, for selecting generating curves for each of its zonotopes (by choices available in a perhaps daunting number of ways) in such a manner that the selected curves approach a limit. Thus, we ask: Given convergent zonotope sequence  $Z_i \rightarrow$  zonoid  $Z$ , how may one determine a convergent sequence of curves  $C_i$  such that  $Z_i \equiv S_{C_i}$ ,  $C_i \rightarrow C$ , and  $S_C \equiv Z$ ?

Many other interesting questions remain. What zonoids do various curves generate—e.g., the "pretty" curves such as a helix? How do other cumulative sweeps behave—the cumulative rotational sweep (CRS), the cumulative mixed sweep (CMS)? What are the properties of each of the cumulative sweeps under constraints of limited memory?

### Acknowledgments

We thank M. Wesley and R. Wolfe for their encouragement and support of this effort, J. Meyer and F. Gracer for assisting with underlying algorithms, B. Watt for rendering the objects shown in Figure 2, and M. O'Connor and J. Rossignac for helpful discussions of theoretical aspects.

### References

1. G. M. Koppelman and M. A. Wesley, "OYSTER: A Study of Integrated Circuits as Three-Dimensional Structures," *IBM J. Res. Develop.* **27**, No. 2, 149–163 (1983).
2. M. A. Wesley, T. Lozano-Pérez, L. I. Lieberman, M. A. Lavin, and D. D. Grossman, "A Geometric Modeling System for Automated Mechanical Assembly," *IBM J. Res. Develop.* **24**, 64–74 (1980).
3. R. N. Wolfe, M. A. Wesley, J. C. Kyle, Jr., F. Gracer, and W. J. Fitzgerald, "Solid Modeling for Production Design," *IBM J. Res. Develop.* **31**, No. 3, 277–295 (May 1987, this issue).
4. J. R. Rossignac and A. A. G. Requicha, "Offsetting Operations in Solid Modelling," *Comput. Aided Geom. Design* **3**, 129–148 (1986).
5. J. R. Rossignac, "Blending and Offsetting Solid Models," *TM 54 Production Automation Project* (also Ph.D. dissertation), University of Rochester, New York, June 1985.
6. R. T. Farouki, "The Approximation of Non-Degenerate Offset Surfaces," *Comput. Aided Geom. Design* **3**, 15–43 (1985).
7. R. T. Farouki, "Exact Offset Procedures for Simple Solids," *Comput. Aided Geom. Design* **2**, 257–279 (1985).
8. B. Chazelle, "Convex Partitions of Polyhedra: A Lower Bound and Worst-Case Optimal Algorithm," *SIAM J. Comput.* **13**, No. 3, 488–507 (1984).
9. T. Lozano-Pérez and M. Wesley, "An Algorithm for Planning Collision-Free Paths Among Polyhedral Obstacles," *Commun. ACM* **22**, 560–570 (1979).
10. J. U. Korein, "A Geometric Investigation of Reach," ACM Distinguished Dissertation, MIT Press, Cambridge, MA, 1985.
11. T. Lozano-Pérez, "Spatial Planning: Configuration Space Approach," *IEEE Trans. Computers* **C-32**, No. 2, 108–120 (1983).
12. W. J. Dally, W. Donath, and D. Ling, "Fast Convolution Operation for Contact Verification in Integrated Circuits," *IBM Tech. Disclosure Bull.* **28**, No. 12, 5588–5594 (May 1986).
13. H. S. M. Coxeter, *Regular Polytopes*, Dover Publications, New York, 1973, pp. 27–30.
14. E. S. Fedorov, "Elemente der Gestaltenlehre," *Mineralogischesche Obshchestvo Leningrad* **21**, No. 2, 1–279 (1885).
15. B. Grünbaum, *Convex Polytopes*, Interscience Publishers, New York, 1967.
16. P. McMullen, "On Zonotopes," *Trans. Amer. Math. Soc.* **159**, 91–109 (1971).
17. J. Serra, *Image Analysis and Mathematical Morphology*, Academic Press, Inc., New York, 1982.
18. Stanley R. Sternberg, "An Overview of Image Algebra and Related Architectures," *Integrated Technology for Parallel Image Processing*, Academic Press, Inc., New York, 1985, pp. 79–100.
19. R. Schneider and Wolfgang Weil, "Zonoids and Related Topics," *Convexity and Its Applications*, Peter Gruber and Jörg Wills, Eds., Birkhäuser, Cambridge, MA, 1983, pp. 296–317.
20. U. Betke and P. McMullen, "Estimating the Sizes of Convex Bodies from Projections," *J. Lond. Math. Soc.* **27**, No. 2, 525–538 (1983).

Received September 30, 1986; accepted for publication January 28, 1987

**Roger C. Evans** IBM Thomas J. Watson Research Center, P.O. Box 218, Yorktown Heights, New York 10598. Mr. Evans joined IBM as a Research staff member in 1966, after receiving in that year an M.A. in mathematics from the University of California at Berkeley and holding a prior A.B. in history from Yale University, New Haven, Connecticut. He worked first as a member, and later as

manager, of a group of applications programming consultants. In 1972 he joined the automation research project to help develop the first IBM robotics system, later receiving a Research Division Award for his work. Between 1976 and 1982, after a move to the text analysis field, he developed pattern description and detection software and applied these to the automatic analysis of business letters. He rejoined, and has been with, automation research since 1982, working in recent periods on polyhedral modeling of kinematic mechanisms and of wafer fabrication processes.

**George Koppelman** *IBM Thomas J. Watson Research Center, P.O. Box 218, Yorktown Heights, New York 10598.* Mr. Koppelman is a Research staff member in the Manufacturing Research Department at the Thomas J. Watson Research Center. He received his B.S. from the University of Chicago, Illinois, in 1961 and joined IBM at the Research Center in 1965, working in pattern recognition, graphics, and interactive document scanning. Subsequently, for several years, he participated in the Josephson technology program, working on circuit simulations, design automation tools for master image logic chips, and microprocessor design. These interests led to a period in the VLSI design group implementing a custom chip design tool. His current involvement in developing the OYSTER wafer fabrication modeling system grew out of this background in graphics and patterning software combined with simulations of device hardware and processor logic.

**V. T. Rajan** *IBM Thomas J. Watson Research Center, P.O. Box 218, Yorktown Heights, New York 10598.* Dr. Rajan is a Research staff member in the design automation group at the Thomas J. Watson Research Center. He received his B.Sc. from the Indian Institute of Technology, Kharagpur, and his Ph.D. in physics from the University of California, Berkeley, in 1976. Since he joined IBM in 1984, he has worked on the application of computational geometry to geometric modeling. He is interested in efficient algorithms for geometric processing, data structures that support these algorithms, and issues arising from their implementation.